Diffusion Monte Carlo for accurate dissociation energies of 3d transition metal containing molecules

Katharina Doblhoff-Dier,*,† Jörg Meyer,† Philip E. Hoggan,‡ Geert-Jan Kroes,† and Lucas K. Wagner¶

†*Leiden Institute of Chemistry, Gorlaeus Laboratories, Leiden University, Post Office Box 9502, 2300 RA Leiden, Netherlands.*

‡*Institute Pascal, UMR 6602 CNRS, University Blaise Pascal, 4 avenue Blaise Pascal, TSA 60026, CS 60026, 63178 Aubiere Cedex, France*

¶*Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

E-mail: k.doblhoff-dier@umail.leidenuniv.nl

Abstract

Transition metals and transition metal compounds are important to catalysis, to photochemistry and to many superconducting systems. We study the performance of diffusion Monte-Carlo (DMC) applied to transition metal containing dimers (TMCDs) using single-determinant Slater-Jastrow trial wavefunctions and investigate the possible influence of the locality and pseudo-potential errors. We find that the locality approximation can introduce non-systematic errors of up to several tens of kcal/mol in the absolute energy of Cu and CuH if Ar or Mg core pseudo-potentials (PPs) are used for the 3d transition metal atoms. Even for energy differences such as binding energies, errors due to the locality approximation can be problematic if chemical accuracy

is sought. The use of the Ne-core PPs developed by Burkatzki et al. (J. Chem. Phys., 129, 164115, 2008), the use of linear energy minimization rather than unreweighted variance minimization for the optimization of the Jastrow function and the use of large Jastrow parametrizations reduce the locality errors. In the second section of the paper, we study the general performance of DMC for 3d TMCDs using a database of binding energies of twenty TMCDs, for which comparatively accurate experimental data is available. Comparing our DMC results to these data, in our results that compare best with experiment, we find a mean unsigned error (MUE) of 4.5 kcal/mol. This compares well with the achievable accuracy in CCSDT(2)$_{\mathrm{Q}}$ (MUE = 4.6 kcal/mol) and the best all-electron DFT results (MUE = 4.5 kcal/mol) for the same set of systems (Truhlar and co., J. Chem. Theory and Comp., 234105, 2015). The mean errors in DMC depend less on the exchange-correlation functionals used to generate the trial wavefunction than the corresponding mean errors in the underlying DFT calculations. Furthermore, the QMC results obtained for each molecule individually vary less with the functionals used. These observations are relevant for systems such as molecules interacting with transition metal surfaces where the DFT functionals performing best for molecules (hybrids) do not yield improvements in DFT. Overall, the results presented in this paper yield important guidelines for both the assessment of the achievable accuracy with DMC and the design of DMC calculations for systems including transition metal atoms.

1 Introduction

Transition metals and their compounds constitute interesting chemical systems. They are omnipresent in catalytic reactions due to their numerous oxidation states and their ability to adsorb molecules on their surfaces; furthermore, they often have interesting magnetic and attractive transport properties (electronic transport, superconductivity, thermal transport). Unfortunately, the energetic properties of transition metal compounds are difficult to calculate due to their complicated electronic structure and the presence of strong correlation.

Although the ultimate goal may be to study catalytic reactions, superconductivity and other bulk properties, small transition metal containing molecules are ideal test-cases for electronic structure methods, since the small number of electrons involved still allows comparatively complex calculations such as large scale CASPT2 or CI calculations to be performed. Furthermore, the correct description of the interaction with the transition metals, which can be tested easily in these small systems, will also be of crucial importance when describing bulk phenomena and interactions with transition metal surfaces. Consequently, various studies on transition metal containing dimers (TMCDs) have been published. The methods used include high level quantum chemical methods such as CASPT2, CI and CCSD(T) (e.g. Ref. 1–3), and also DFT or DFT+U (e.g. Ref. 4–7). Although some DFT functionals give astonishingly accurate results on average[7], the strong correlation leads to a large dependence on the exchange-correlation (XC) functional (see for example Tab. 1). In fact, even the more advanced methods and in particular the "gold standard" of quantum chemistry, CCSD(T), have problems describing the interactions in TMCDs accurately: mean unsigned errors of 4.5 kcal/mol for CCSD(T) and similar values for $\mathrm{CCSDT(2)_Q}$ have been found by Truhlar and co. in Ref. 7 (see again Tab. 1).

More recently, also quantum Monte Carlo (QMC) methods have been applied to transition metal atoms[11,12], TMCDs[10,13–17], clusters[18] and to solids[19,20] including transition metals. The interest in applying QMC methods, such as reptation Monte Carlo and diffusion Monte Carlo (DMC), to these systems is sparked by the fact that QMC methods are able to treat correlation effects explicitly rather than using approximate density functionals, while scaling as $\mathcal{O}(N^3)+c\mathcal{O}(N^4)$, where c is a small constant and N is the number of particles, if a constant statistical error-bar in the total energy is sought and localized basis functions are used[21,22]. This scaling with system size allows the treatment of large molecular complexes and solids such as transition metal oxides. This would be impossible to achieve with CCSD(T) and other explicitly correlated quantum chemistry methods.

In spite of the comparative wealth of DMC studies applied to transition metal containing

Table 1: Comparison of the mean unsigned error (MUE) of binding energies for a database of 20 3d TMCDs obtained with different methods. (Data taken from Ref. 7 and this work.) Basis set definitions: *a*: def2-TZVP, minimally augmented on the ligands; *b*: aug-cc-pwCVTZ-DK[8], aug-cc-pVTZ-DK for ligands; *c* aug-cc-pwCVTZ-DK[8] (see Ref. 7 for details); BFD-PP: small core pseudo-potentials by Burkatzki et al.[9,10] DMC(B97-1) denotes a DMC calculation based on a Slater-Jastrow trial wavefunction generated from a B97-1 Slater function. The last result for DMC(B3LYP) with the note "Padé func." was obtained using a Jastrow factor parametrized by Padé functions in QWALK, while all other DMC calculations were obtained with polynomial Jastrow functions in CASINO.

method	details	MUE [kcal/mol]	Ref.
B97-1	all e^-, non-rel., bas. *a*	5.1	Ref. 7
B3LYP	all e^-, non-rel., bas. *a*	6.1	
M05	all e^-, non-rel., bas. *a*	7.1	
B97-1	all e^-, scalar rel., bas. *b*,	4.5	
CCSD(T)	all e^-, scalar rel., bas. *c*,	4.5	
$CCSDT(2)_Q$	all e^-, scalar rel., bas. *b*	4.6	
B97-1	BFD-PP, TZV	6.3	this work
B3LYP	BFD-PP, TZV	7.9	
DMC(B97-1)	BFD-PP	5.4	
DMC(B3LYP)	BFD-PP	5.3	
DMC(B3LYP) Padé func.	BFD-PP	4.5	

dimers (e.g. Refs. 10,13–17), the results do not allow one to judge the overall performance of QMC for TMCDs for several reasons. First, the above references only cover transition metal oxides and transition metal sulfides. Next, the studies all used different computational set-ups (different pseudo-potentials, some of which are not readily available, different trial wavefunction generation methods, different parametrizations of the Jastrow function and different move schemes in DMC). Additionally, all mentioned references compare the DMC energies directly to the experimental energies, taking the scalar relativistic effects included in the pseudo-potentials into account but ignoring the effect of spin-orbit coupling. This effect can, however, be on the order of several kcal/mol in TMCDs. Last but not least, no reliable experimental data exists for comparison for most molecules studied and — as mentioned above — even the comparison to CCSD(T) results is problematic for TMCDs.

In our present work, we therefore aim at reassessing the performance of DMC by studying the bonding in a wide range of TMCDs in a systematic and consistent way. More precisely, we study a database of 20 TMCDs with experimental error bars in the dissociation energy lower than 2 kcal/mol. As mentioned above, such a database has been published by Truhlar and co.[7] and it includes best estimates of the bond length and estimates for the spin-orbit coupling calculated by quantum chemistry methods. On the basis of this database, we address important questions, such as "How well can DMC describe the bonding in transition metal systems when using single-Slater-Jastrow trial wavefunctions?", "Does it perform as well as or better than the best density functionals without depending so strongly on the functional used in the trial wavefunction generation?" and "Can QMC, in spite of its good scaling with system size, compete with CCSD(T)?" Additionally, we address several other questions, which are of specific relevance for DMC calculations including transition metals. Most importantly, we address the performance of different optimization schemes for the trial wavefunctions and the influence of the locality approximation[23,24] — an approximation, which is necessary in DMC calculations when non-local PPs are used (see Sec. 2).

The paper is structured as follows: in Sec. 2, we introduce the methods used and present

the computational details. In Sec. 3, we discuss tests performed for Cu and CuH concerning the choice of pseudo-potentials and optimization method. Sec. 4, focuses on the results obtained for the database of transition metal molecules tested. Sec. 5 comprises the summary and the conclusion.

2 Methods

2.1 Variational Monte Carlo and diffusion Monte Carlo

In the course of this work, we use two quantum Monte Carlo methods, namely variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC). These methods are described in length in, for example, Refs. 25–27. Here, we shall therefore only give a brief introduction.

In VMC, a variational wavefunction

$$\psi_T = \psi_S \cdot e^{J(\{r_{ij}, r_{Ii}, r_{Ij}\})} \tag{1}$$

is constructed by multiplying a Slater determinant, ψ_S, by a so-called Jastrow factor e^J. The Jastrow function J is a function of electron-electron distances r_{ij} and ion-electron distances (r_{Ii} and r_{Ij}) and serves to introduce many particle correlation effects. Since the integrals involved in evaluating the energy of the resulting many-particle correlated wavefunction are high-dimensional, the variational energy $E_V = \frac{\langle\psi_T|\hat{H}|\psi_T\rangle}{\langle\psi_T|\psi_T\rangle}$ is evaluated stochastically using Monte Carlo sampling and the Metropolis algorithm.

The Jastrow function J is parametrized, generally by several tens to a few hundreds of parameters, and is optimized either by minimizing the variance of the local energies $E_L(\mathbf{R}) = \frac{\hat{H}\psi(\mathbf{R})}{\psi(\mathbf{R})}$ ("variance optimization") or by minimizing the variational energy ("energy optimization") for a finite set of configurations $\{\mathbf{R} = (\mathbf{r}_1, \ldots, \mathbf{r}_N)\}$ of the electronic coordinates $\mathbf{r}_i$. Traditionally, only variance minimization was used, since stable and efficient algorithms allowing the optimization of the energy were only derived in the very late 1990s

and the 2000s (e.g. Ref. 28; we refer the reader to Ref. 29 for a more detailed discussion of the development of optimization schemes).

In DMC, an initial wavefunction $\psi(\tau = 0) = \psi_T$ is propagated in imaginary time τ according to the imaginary time Schrödinger equation $-\frac{\mathrm{d}\psi}{\mathrm{d}\tau} = \hat{H}\psi$ (atomic units are used). Subject to the fixed-node constraint, in which the nodes of the wavefunction $\psi(\tau)$ are fixed to those of the trial wavefunction ψ_T, this allows the fixed-node ground-state $\psi_0^{\mathrm{FN}} = \psi(\tau \to \infty)$ to be projected out and the fixed-node energy E to be calculated. The projection itself is carried out using stochastic approaches and moving a set of configurations according to a drift-diffusion and rate equation resulting from the imaginary time Schrödinger equation at a finite time-step $\Delta\tau$.

If non-local pseudo-potentials are used, an additional approximation is necessary in DMC calculations. The use of PPs is very common in DMC due to the unfavourable scaling of the computational cost[30,31] with the nuclear charge Z ($\mathcal{O}(Z^{5.5})$ to $\mathcal{O}(Z^{6.5})$). The evaluation of the non-local potential, however, requires approximations. The so-called locality approximation consists in evaluating the non-local part of the pseudo-potential on the trial wavefunction ψ_T, instead of on ψ. Traditionally[23,24], this localized potential was added to the local potential. This has the disadvantage of giving DMC energies that may or may not satisfy the variational principle (i.e., which can be lower than the true fixed-node energy). There is, however, a slightly different way of treating the non-local pseudo-potential, which is known as the "T-move" scheme. This scheme, in which the non-local potential is not fully localized, re-establishes the variational principle[32–34].

2.2 Computational details

Four our comprehensive study on TMCDs, we use the same database of TMCDs as Ref. 7. For completeness, the 20 molecules present in this database are listed in Tab. 2. Their experimentally measured bond dissociation energies D_e (corrected for zero-point energy) lie between 21.6 kcal/mol (ZnH) and 151.0 kcal/mol (VO) with maximum experimental uncer-

tainties of 2.0 kcal/mol.

Table 2: List of the 3d TMCDs studied together with their experimental bond dissociation energy $D_e^{\text{exp.}}$ (corrected for zero-point energy) and the experimental uncertainty thereof. Data taken from Ref. 7.

	$D_e^{\text{exp.}}$ [kcal/mol]	uncert. [kcal/mol]		$D_e^{\text{exp.}}$ [kcal/mol]	uncert. [kcal/mol]
TiCl	100.8	2.0	FeCl	78.5	1.6
VH	51.4	1.6	CoH	45.5	1.2
VO	151.0	2.0	CoCl	80.5	1.6
VCl	101.9	2.0	NiCl	88.0	1.0
CrH	46.8	1.6	CuH	62.6	1.4
CrO	104.7	1.2	CuCl	87.7	0.4
CrCl	90.1	1.6	ZnH	21.6	0.5
MnS	70.5	2.0	ZnO	37.9	0.9
MnCl	80.7	1.6	ZnS	34.3	1.0
FeH	36.9	0.8	ZnCl	53.5	1.0

Although all-electron DMC calculations may be achievable for TMCDs, the computational cost associated with all-electron computations of crystals and solids including heavy atoms renders such large-scale calculations impossible with current computational resources. We therefore chose to apply pseudo-potentials (PP) even for the dimer calculations. The choice of pseudo-potential is of special relevance in DMC, since the PP does not only determine the physical description of the system in terms of core-valence correlation but can also have a large impact on the errors in DMC due to the locality approximation[23,24,33,34]. For the PP tests in Sec. 3, we use several different PPs. More precisely, we use three different Ar core PPs: 1.) an Ar core Troullier-Martins-type[35] PBE PP from the standard PP set in Abinit[36] with an s local channel (i.e. $l_{\text{loc}} = 0$) , 2.) an in-house developed Ar core Troullier-Martins-type PBE PP generated with the fhi98PP code[37] with local channel f (i.e. $l_{\text{loc}} = 3$) and 3.) a Trail-Needs Ar core Dirac-Fock average relativistic effective PP[38,39] ($l_{\text{loc}} = 2$ for QMC, $l_{\text{loc}} = 0$ for DFT to avoid ghost states) . Furthermore, we used two PPs with smaller core sizes, namely a Mg core Trail-Needs Dirac-Fock average relativistic effective PP[39] ($l_{\text{loc}} = 2$ for QMC, $l_{\text{loc}} = 0$ for DFT to avoid ghost states), and the Ne core PPs generated for the use in DMC by Burkatzki, Filippi and Dolg[9,10] (denoted by BFD-PPs

in the following) ($l_{\rm loc} = 2$). Based on these test calculations for Cu and CuH (see Sec. 3), we chose to apply the BFD-PPs to all atoms except hydrogen[a] in the database calculations. These PPs do not only allow for small locality errors (see Sec. 3), their Ne core for the 3d metals also allows the correlation of the semi-core electrons. This has been shown to be important for transition metals (e.g. Refs. 8,41–47). We note that the BFD-PPs are of course not the only possible choice of small core PPs that can be used in diffusion Monte Carlo. The PPs from the Stuttgart-Köln table[48], for example, can be adapted to allow their use in diffusion Monte Carlo by smoothing the divergence at the center of the PP, as has been done in the past (e.g. Refs. 13,14,44).

The Slater part ψ_S of the trial wavefunction is generated with the program Gamess-US[49] and a version of that code[50] that allows direct interfacing to CASINO[51]. For the calculations in Sec. 3 Gaussian09 and pwscf[52] are used. All single determinant wavefunctions are calculated without symmetry constraints, allowing convergence to the lowest energy state. The multiplicity is set to the value corresponding to the ground states listed in Ref. 7. The initial guess used in the wavefunction generation was chosen to be compatible with the experimental ground state symmetry, also listed in Ref. 7. We note that this may not necessarily lead to convergence to the global minimum in DFT. On the other hand, this approach is reproducible and will lead to trial wavefunctions that are most closely related to the experimental structure, which is an advantage in DMC. Multi-determinant calculations were based on a second-order configuration interaction (CI) calculation using the molecular orbitals from restricted open shell B3LYP calculations as in Ref. 16. The CI space included a full CI in the partially occupied degenerate subspace and singles and doubles excitations to unoccupied orbitals. As basis sets, the following sets were used: a.) the triple zeta version of the basis given in Refs. 9,10 without g-functions but augmented by the augmentation set of the Dunning correlation consistent basis to ensure good basis set coverage even at the relatively long bond distances of some of the dimers (single Slater wavefunctions), b.) the

[a]We note parenthetically that a hydrogen PP[40] was used for the PP test calculations

same triple zeta basis with g-functions but without augmentation (CI calculations) and c.) a plane wave basis for calculations including the Ar core PPs (plane wave cutoff: 80 Hartree) and the Ne core PP (plane wave cutoff: 160 Hartree). For the PP tests in Sec. 3, the Slater wavefunction is constructed using the method which was used to generate the PP, i.e. with Hartree-Fock calculations for the BFD-PP and the Trail-Needs PPs, and with PBE calculations for the Troullier-Martins-type PBE PPs.

The QMC part of the calculations was mainly performed using the CASINO program[51]. In some cases, QWALK[53] was used additionally in order to take advantage of slightly differing capabilities. In all calculations, the Jastrow function J is taken as a sum of one-, two- and three-body terms ($\chi(r_{iI})$, $u(r_{ij})$ and $f(r_{ij}, r_{iI}, r_{jI})$)

$$J(\{\mathbf{r}_i\}, \{\mathbf{r}_I\}) = \sum_{I,i} \chi(r_{iI}) + \sum_{i,j\neq i} u(r_{ij}) + \sum_{I,i,j\neq i} f_I(r_{iI}, r_{jI}, r_{ij}). \tag{2}$$

In the calculations performed with the CASINO program package, the functions $\chi(r_{iI})$ and $u(r_{ij})$ are given by a polynomial of degree N multiplied by a cutoff-function for each of their variables x:

$$P_N(x) \cdot (x - L_x)^3 \cdot \Theta(L_x - x). \tag{3}$$

Here, $P_N(x)$ is a polynomial of degree N, L_x denotes the cutoff-length with respect to variable x and Θ is the Heavyside Theta function, which is zero for negative values and one otherwise[54]. The three-body function $f(r_{ij}, r_{iI}, r_{jI})$ is given equivalently as a polynomial in the three independent variables multiplied with a cutoff-function in r_{iI} and r_{jI}. The cutoff-length L_x and the polynomial coefficients are optimizable parameters. Different parameters are used for spin up and spin down electrons in the χ term and for the different spin pairs in the u and the f term. The $u(r_{ij})$ term is restricted to satisfy the electron-electron cusp conditions. Unless explicitly stated, the polynomial degree of the u and the χ terms is 4 and the polynomial degrees in r_{ij}, r_{iI} and r_{jI} in the f term is 2. In the PP tests (Sec. 3), we use several different Jastrow parametrizations, which we discern by the number of parameters

used. The Jastrow function with 31 optimizable parameters contains only one and two body terms with polynomial orders of 4. The Jastrow functions with 81 parameters additionally contain three body terms with polynomial order 2. Finally, the largest parametrization used (256 parameters) uses polynomial orders of 8, 8 and 4 for the u, the χ and the f term, respectively.

Calculations performed with the QWALK program package use a different form of the Jastrow function. In this case, the one-, two- and three-body terms are parametrized using functions $P_k(x)$, which have the form of a Padé approximant

$$P_k(x;\beta_0) = \frac{1 - z(x/L_x)}{1 + \beta_k \cdot z(x/L_x)}. \tag{4}$$

In this expression, $\beta_k = \exp(\beta_0 + 1.6k) - 1$ and $z(y) = y^2(6 - 8y + 3y^2)$. The one-, and two-body terms are each parametrized by a function g_1

$$g_1 = \sum_k c_k P_k(x;\beta_0), \tag{5}$$

where the index k runs from zero to 3 if x denotes an electron-electron variable and from zero to 4 if x denotes an electron-nucleus variable[55]. The optimizable parameters are the c_k and β_0. These parameters differ for the χ-function (c_k^χ and β_0^χ) and the u-function (c_k^u and β_0^u). The three-body term is given as

$$f = \sum_{klm} c_{klm}^f \left[P_k(r_{iI};\beta_0^\chi)P_l(r_{jI};\beta_0^\chi) + P_l(r_{iI};\beta_0^\chi)P_k(r_{jI};\beta_0^\chi)\right] P_m(r_{ij};\beta_0^u). \tag{6}$$

Several of the c_{klm}^f are restricted to zero and only a total of 12 terms is used for the f-function. Since the Padé functions cannot satisfy the cusp conditions, these are taken care of by a function[55]

$$\frac{c \cdot L_x \cdot \tilde{z}(x/L_x)}{1 + \gamma \cdot \tilde{z}(x/L_x)}, \tag{7}$$

where x is the distance between the two electrons, c is 1/4 or 1/2, for same-spin and opposite-

spin electrons respectively, and $\tilde{z}(y) = y - y^2 + y^3/3$. The parameter γ is optimizable.

In all calculations performed with CASINO, the Jastrow factor J is optimized first by minimizing the variance without reweighting the summands according to the wavefunction change at the walker location ("variance optimization")[56,57] and thereafter by minimizing the energy of the trial wavefunction $\psi_T = \psi_S \cdot e^J$ ("energy optimization") via a linear method[28,54,58]. For the initial variance optimization, 50 000 configurations are used in each of the 5 optimization cycles. Several optimization cycles are used to allow the configurations to adapt to the optimized function. In the energy minimization 100 000 configurations are used in each of the five optimization cycles to allow the linearized method to converge. Tests on several of the atoms and molecules in the database showed that an additional energy optimization cycle with twice as many configurations slightly lowered the variational Monte Carlo (VMC) energy in some cases, but no relevant influence on the DMC energies was found. The optimization strategy is thus considered to give converged results (the observed differences were within the statistical error bars of less than 0.5 kcal/mol). The general procedure for the PP test calculations is similar, but the exact parameters may deviate.

In the calculations performed with QWALK, the procedure is similar, with the only difference that the cutoff-length L_x is kept fixed at 7.5 a u and that only one optimization cycle (i.e. one set of configurations) is used for both variance and energy optimization. The number of configurations used in QWALK also differs from that used in CASINO: For variance minimization close to 10 000 configurations were used, for energy optimization up to about 300 000 configurations were used, depending on the convergence.

For some of the calculations in Sec. 4.3.3, also a backflow transformation[59] $\zeta(\{\mathbf{r}_i\})$ is applied to the coordinates $\mathbf{r}_i$ of the Slater determinant (in addition to multiplying with a Jastrow factor):

$$\mathbf{x}_i = \mathbf{r}_i + \zeta(\{\mathbf{r}_i\}). \tag{8}$$

Such a transformation allows the nodes of the trial wavefunction to be shifted and to introduce further correlation to the trial wavefunction. Similarly to the Jastrow function, $\zeta(\{\mathbf{r}_i\})$

is expanded in one-, two- and three-body terms, each described by a polynomial function with a cutoff function. The polynomial orders used are $N = 6$ for the one- and two-body terms and $N = 2$ for the tree-body terms. The parameters in the backflow transformation are optimized in an additional energy optimization step, in which both the parameters of the backflow function and the parameters of the Jastrow function are allowed to vary.

The optimized trial functions are subsequently used in diffusion Monte Carlo (DMC)[60] calculations, again using the software package CASINO[51] or QWALK[53]. Unless otherwise stated, all energies are computed using the "T-move" scheme[33,34] in order to re-establish the variational principle for the DMC results in spite of the use of non-local pseudo-potentials. In the DMC calculations performed in CASINO, a target number of 7680 configurations is used. After an equilibration length of 850, 1600 or 2000 steps (depending on the time step), the DMC data is accumulated for statistical analysis. The equilibration length used is probably not optimal in terms of computer time for most calculations. Shorter equilibration length could have been chosen thereby reducing the computational effort, but the constant equilibration length allowed straightforward automatic input generation and does not make up a large amount of computing time for most calculations. On the other hand, the equilibration length is checked to be long enough, by comparing the results with results of using a longer equilibration time. To obtain statistical error bars on the data, the DMC data accumulated after equilibration is re-blocked[61] using the automatic re-blocking facility of CASINO.

In QWALK, 9600 configurations are used. Instead of performing a reblocking analysis, the energies are averaged over subsequent blocks of 1 a.u. in the imaginary time propagation. The equilibration length is automatically determined by detecting the blocks at the beginning of the time series that lie outside the error bars. In order to minimize time-step errors, we perform calculations at $\tau = 0.008$ au, $\tau = 0.005$ au and $\tau = 0.002$ au. Subsequently, a linear extrapolation is performed for $\tau \to 0$. We note that the choice of time-step is not the most efficient in terms of computational cost[62], but it allows us to check the validity of the linear extrapolation by comparing the results with results from an extrapolation using only the

two shorter time steps in an extrapolation. The parameters used in the PP tests are similar, but the exact values used may differ.

The bond dissociation energies D_e are calculated from the zero-time-step extrapolated DMC energies E as

$$D_e(MX) = E(M) + E_{\mathrm{SO}}(M) + E(X) + E_{\mathrm{SO}}(X) - E(MX) - E_{\mathrm{SO}}(MX), \qquad (9)$$

where M denotes the transition metal atom and X is the other atom (H, O, S or Cl). The equilibrium geometry of the molecules MX is taken from Ref. 7, as are the spin orbit corrections E_{SO}, which were calculated at the MCSCF level. Scalar relativistic corrections for the core electrons are already taken care of in the PPs. Note that the values of D_e given in Tab. 2 were corrected for zero-point energy by Truhlar and co.[7] using DFT calculations. All statistical error bars obtained for the final dissociation energy lie below 0.4 kcal/mol.

Finally we note that a spin-orbit-correction using MCSCF results may not be possible for larger systems than addressed here due to the unfavourable scaling of MCSCF. For large systems, a spin-orbit correction using relativistic TDDFT[63] may be an option. Furthermore, depending on the system under investigation, spin-orbit corrections may not be necessary, as is also in the case for some of the systems addressed here ($\mathrm{CuH} \rightarrow \mathrm{Cu} + \mathrm{H}$).

3 Pseudo-potential tests and the locality approximation

As mentioned above, the use of PPs is favorable from a computational point of view, but their use is linked to approximations. First of all, the PP gives only an approximate description of the real, all-electron system. Although it has been shown that the use of PPs does not degrade the DMC results for the G2 set of molecules[64,65], the PP error may become more important in more complicated systems, especially if very high accuracy is sought in QMC

calculations (e.g. Refs. 66,67). Apart from these frozen core induced shortcomings of the PPs, the locality error is another source of error in DMC calculations. In the following, we focus on this type of error.

We analyze this error using the Cu atom and the CuH molecule as examples. We start our discussion by demonstrating that the common assumption of a negligible error due to the locality approximation is not generally true for single determinant trial wavefunctions. Then, we discuss how to treat systems in which the errors due to the locality approximation tend to be large. Finally we study the performance of pseudo-potentials with different core sizes in DMC and anlayze error cancellation effects for the CuH binding energy.

3.1 Absolute locality errors in large core PPs

Using an Ar core Troullier-Martins-type[35] PBE PP from the standard PP set in abinit[36] for the Cu-Atom, we found DMC total energies that differed by nearly 80 kcal/mol (3.4 eV) depending on whether variance or energy optimization was used to generate the Jastrow function (see Fig. 1). In calculations which use local potentials only (i.e. no locality approximation), the fixed-node approximation remains the only fundamental approximation in DMC for finite systems and the energies obtained in DMC should, in principle (after converging the population and timestep bias), not depend on the Jastrow function[51], since the latter does not change the nodes. In this case, the DMC results would be independent of the method used to optimize the parameters in the Jastrow function, in spite of the fact that different optimization methods may converge to different Jastrow parametrizations due to the limited flexibility of the parametrization and due to the possibility to get stuck in local minima. The dependence of the DMC energy on the Jastrow function can thus only be ascribed to the locality approximation. Similar results were also found for the other Ar core PPs given in Sec. 2.2.

A possible explanation for the large locality errors might be a failure of the Kleinman-Bylander representation: plane-wave atom/molecule-in-a-box calculations and the Kleinman-

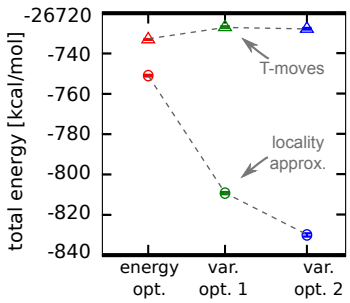

Figure 1: Cu atom absolute DMC energies obtained using a Jastrow factor from energy optimization or variance optimization for an Ar core PP (see text for details). "var. opt. 1" and "var. opt. 2" denote two variance optimization cycles with different sets of configurations used in the optimization. Circles: results using the locality approximation as introduced in Refs. 23,24; triangles: results using T-moves.

Bylander form of the PP were used to generate the Slater-wavefunction for all the above calculations. However, bypassing the atom-in-a-box plane wave calculation entirely by constructing the trial wavefunction based on the pseudo-orbitals obtained from the PP generator fhi98PP, we could show unambiguously that it is not the Kleinman-Bylander representation of the PP which leads to the large locality error.

We therefore conclude that the use of these large core PPs for heavy atoms such as Cu can lead to considerable locality errors and, as shown in the following, care has to be taken in calculations involving such PPs.

3.2 Handling locality errors occurring for the large core PPs used

In principle, a large locality error does not necessarily pose a problem in practical calculations, since the error may largely cancel out when calculating energy differences. This necessitates though, that the results obtained in DMC are stable with respect to stochastic fluctuations in the wavefunction optimization. However, we found the DMC results to vary several tens of kcal/mol when using variance optimization with the results depending strongly on the exact Jastrow function resulting from the optimization (see Fig. 1), which may lead to problems with error cancellation as explained below.

A large enough number of configurations used in the optimization should in principle

allow one to obtain a stable convergence to a minimum, and thus a stable Jastrow function and stable DMC energies. With variance optimization, however, we found that even very large numbers of configurations do not stabilize the Jastrow function sufficiently to obtain stable DMC results. A particularly alarming example was found for a case where we used a very large set of 800 000 configurations in the variance optimization procedure optimizing 49 parameters. Using two different sets of configurations, we found two different Jastrow functions, which gave the same (comparatively low) VMC energy and would thus be expected to give similar DMC results. However, the two trial wavefunctions lead to a DMC energy difference of more than 20 kcal/mol (see Fig. 1). This "instability" in the DMC energies hinders error cancellation when calculating energy differences since we cannot expect "randomly" changing errors to cancel.

For trial wavefunctions resulting from energy optimization, on the other hand, we never observed such drastic changes in energy (although we did observe changes on the order of a few kcal/mol in some cases). We explain the possible superiority of energy minimization over variance minimization as follows: there are indications that the wavefunction itself is better represented when using energy optimization than when using variance optimization. This claim is based on the observations that expectation values other than the energy, which are based on a mixed estimator, can be computed more accurately using energy minimized trial wavefunctions[68,69]. A better trial wavefunction would also decrease the locality error, which is proportional to the difference between the trial wavefunction and the exact eigenstate squared[24]. We therefore suggest energy minimization as the preferable optimization scheme when using pseudo-potentials.

Using the T-move[33,34] scheme instead of normal moves, we made two interesting observations: First, we found that the energies obtained using energy optimized Jastrow functions are generally lower in energy than those found using variance optimization (see Fig. 1 as example). Secondly, we found that the dependence of the DMC energy on the Jastrow function is much reduced. To give an example, we only obtain an energy difference of about

5 kcal/mol between the above mentioned energy and variance optimized results, instead of the 80 kcal/mol difference when using the locality approximation from Refs. 23,24 (see Fig. 1). The difference in DMC energy for the two examples of variance minimization was even reduced to 0.8 ± 0.6 kcal/mol. The first observation (lower DMC energies for energy optimized trial wavefunctions) backs our suggestion that energy optimization should be the preferred optimization scheme and the second observation (smaller dependence of DMC energies on the Jastrow function) is of interest for obtaining more stable results.

From our experience, we therefore conclude that, contrary to variance minimization, energy minimization and T-moves allow us to generate stable DMC results to be obtained even for large core PPs. Energy optimization, and T-moves should therefore be used if non-local potentials are present.

3.3 Locality errors with different core sized PPs

To get a quantitative estimate of the locality error with different PPs, we studied four different PPs with different core sizes and local channels in more detail. More precisely we compare results for the two Ar core PBE PPs (denoted by a and b in the following), the Mg core PP (c) and the Ne core PP (d) (all PPs are detailed in Sec. 2.2). Focussing on potential error cancellation for binding energies, we now turn to CuH.

Figure 2 shows the total DMC energy for CuH resulting from trial wavefunctions with different sized Jastrow-parametrizations. With increasing flexibility of the Jastrow function (i.e. increasing accuracy of the trial wavefunction), the DMC energies obtained with T-moves decrease, as expected. With the locality approximation (i.e. no T-moves), the energy converges from below in all cases. This behavior allows us to estimate the locality error from the difference between the result using T-moves and that using the locality approximation. This difference (measured at the largest Jastrow parametrization) is approximately 13 kcal/mol for the Ar-core PPs, 6 kcal/mol for the Mg-core PP and 3 kcal/mol for the Ne-core PP. If we want to obtain chemically accurate binding energies (errors < 1 kcal/mol) in spite of the

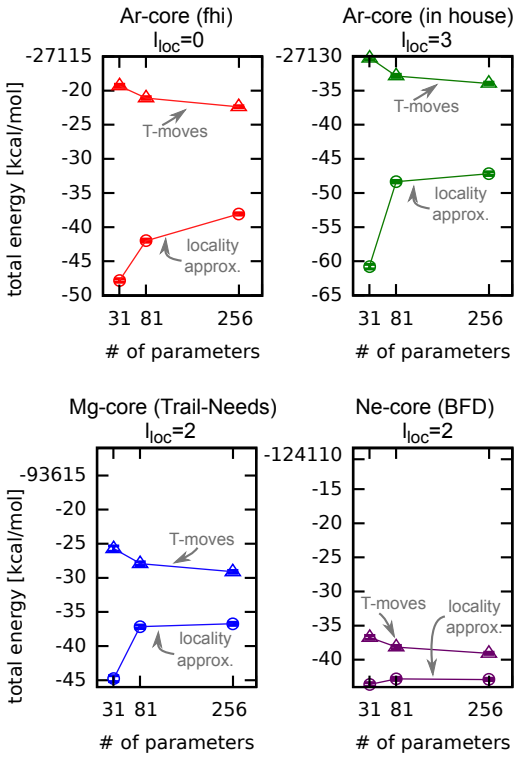

Figure 2: DMC absolute energy for CuH using different PPs and different sizes of Jastrow parametrizations as expressed by the number of Jastrow function parameters on the x-axis (parameters optimized using energy optimization). Triangles denote results obtained with T-moves, circles denote results using the locality approximation as introduced in Refs. 23,24. (For details on the parametrizations see Sec. 2.2).

.

presence of locality errors, we have to rely on error cancellation in the calculation of the corresponding energy differences.

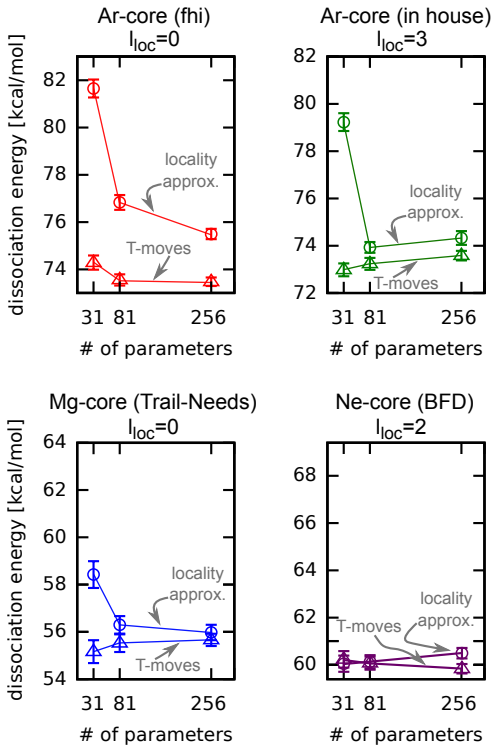

Figure 3: DMC dissociation energy of CuH (see Fig. 2 for details). Note that the y-axis is shifted between the different plots, but that the scale is constant throughout the 4 plots.

To test the effect of error cancellation, Figure 3 shows the CuH dissociation energies obtained with and without T-moves. In the case of the Ar-core PPs, for the smallest Jastrow parametrization (which did not include three-body terms), the CuH dissociation energy changes by nearly 8 kcal/mol depending on whether T-moves are used or not. Furthermore, if no T-moves are used, the dissociation energy changes by more than 6 kcal/mol from the smallest to the largest Jastrow parametrization. This suggests that the influence of the locality error is too large to obtain precise results for the large core PPs with small Jastrow parametrizations and without T-moves. If T-moves are used, the dependence on the Jastrow parametrization is weaker. Large Jastrow parametrizations, including three-body terms are also clearly favourable. For the in-house Ar core PP, the dissociation energies then vary only by slightly more than 1 kcal/mol and for the Mg core PP the observed variations are

then even lower than 1 kcal/mol. If the Ne-core PP is used, independently of the Jastrow parametrization used and of whether T-moves are used or not, the dissociation energy does not change by more than 1 kcal/mol. In all these cases, error cancellation leads to changes in the binding energies (i.e., energy differences) that are much smaller than the changes in the absolute energies.

Nevertheless, the PPs with different core sizes result in strongly different values for the dissociation energy (Ar-core $\sim$ 74 kcal/mol; Mg-core $\sim$ 56 kcal/mol; Ne-core $\sim$ 60 kcal/mol; experiment 62.6 kcal/mol[7]). Although this difference may still be influenced by the locality approximation, the fact that the energy differences become stable for large enough Jastrow parametrizations suggest that the "physical" deficiencies (transferability problems) of the PP and the fixed-node approximation rather than the locality error cause the deviation from the experiment. Large core PPs can be expected to result in sizeable errors, since core-valence correlation has been shown to be important even for atomic properties of 3d transition metals as early as the 1980's (e.g. Refs. 8,41–47). However, in systems where (some) transition metal atoms act as spectators and are not directly involved in the process of interest, the errors due to the missing core-valence correlation may cancel out to a high extent allowing for the use of large core PPs, but only if the "random" errors stemming from the locality approximation are not too large.

Although we have shown in Fig. 3 that the use of large enough Jastrow parametrizations and the use of T-moves in principle allows using the large(r) core PPs in DMC, we choose the Ne-core BFD-PPs for the remaining work for several reasons: 1.) we expect semi-core valence correlation to be of considerable importance in some of the transition metal atoms and transition metal containing molecules (see above). This is also confirmed by the CuH results shown above, for which the BFD-PP result is closest to the experiment, indicating that the other PPs suffer from missing core-valence correlation; 2.) the results for the BFD-PP depend the least on the choice of Jastrow parametrization.

With this, we now turn to the more general assessment of the achievable accuracy of

DMC for TMCDs and discuss our results for the database of 20 transition metal containing molecules.

4 Database results

4.1 DFT results

We start our discussion of the database results with the DFT and Hartree-Fock (HF) results. This discussion is important for the DMC results, since the DFT and HF results will determine the quality of the wavefunction nodes and may in turn be influenced by the quality of the pseudo-potentials and basis set used. Unlike DMC results, which follow the variational principle if T-moves are used, the absolute energies for different XC-functionals in DFT cannot be compared against each other since they to not follow the variational principle. To judge the quality of these results, we therefore look at the deviations of the dissociation energies from the experimental values. For UHF, PBE, B3LYP and B97-1, we obtain a mean unsigned error (MUE) in the dissociation energy of 35.2 kcal/mol, 8.4 kcal/mol, 7.9 kcal/mol and 6.3 kcal/mol, respectively (see Tab. 4). As expected, the Hartree-Fock results are considerably worse than the DFT results and within the DFT results there is a considerable spread: As one might expect, the hybrid functionals B3LYP and B97-1 perform better than PBE, which is based on the generalized gradient approximation. The results for B3LYP and B97-1 are, however, often considerably different for a specific molecule, in spite of the fact that they both show comparatively small mean errors (see Fig. 4a). This makes it difficult to decide a priori which of the two functionals is better for a particular system.

To asses the quality of the PPs and the basis set used, we also compare our results with results from Ref. 7, where a minimally augmented def2-TZVP basis and all-electron calculations were used. Neglecting scalar relativistic effects, Ref. 7 gives a MUE of 6.1 kcal/mol for B3LYP and 5.1 kcal/mol for B97-1 for consistently optimized bond-lengths (see Tab. 3). Including scalar relativistic effects and using the reference geometries, a MUE of 4.9 kcal/mol

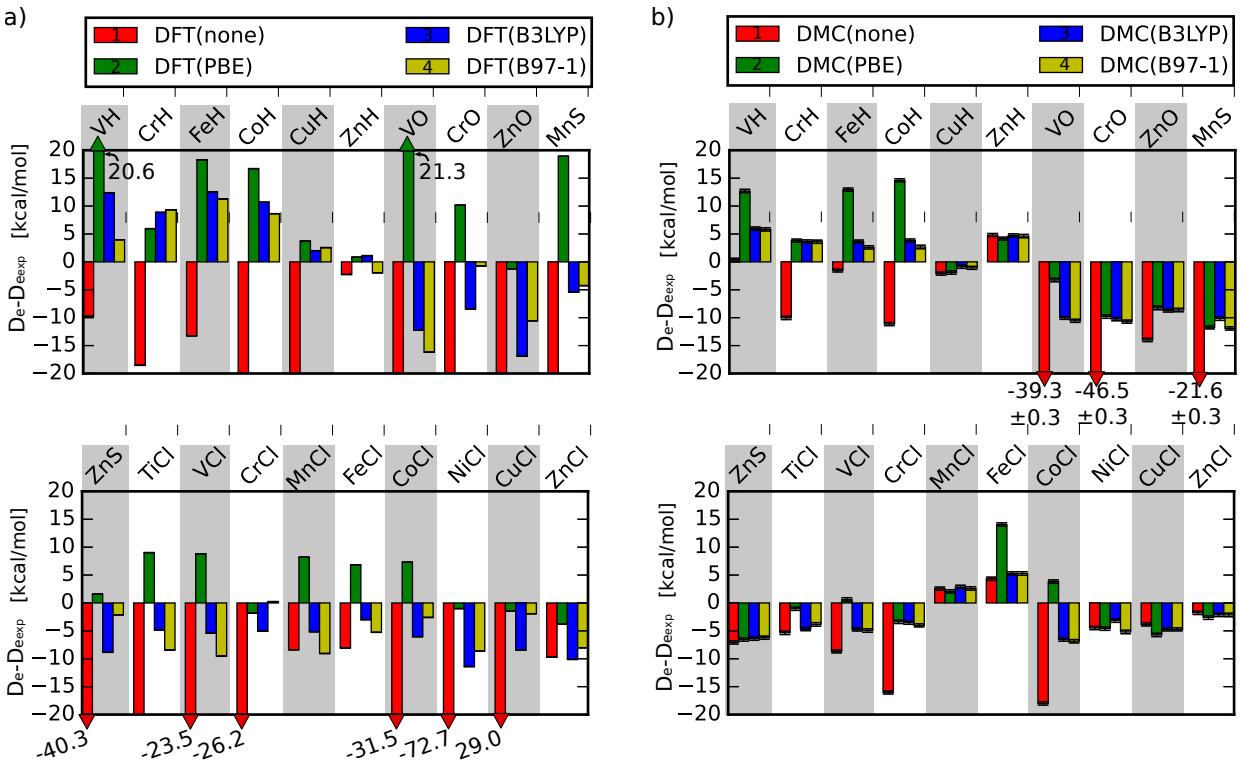

Figure 4: Difference of dissociation energies calculated with DFT (panel a) and DMC (panel b) corrected for spin-orbit coupling with the experimental values. Negative errors indicate underbinding. Errors that are larger than the scale chosen, are given as numeric values. The error-bars (which are barely visible on this scale) include only the uncertainties due to the stochastic nature of DMC and not the experimental uncertainties.

is found for B97-1. The errors we obtain are thus about 1.2 kcal/mol to 1.8 kcal/mol larger. As mentioned in Sec. 2.2, the states we obtain may not in all cases correspond to the lowest possible DFT energy. We reanalyzed our database performing a stability analysis in Gaussian09[70] and reoptimizing the orbitals if necessary. This procedure slightly lowers the DFT errors (see Tab. 3), but the MUE we observe are still larger than those in Ref. 7. We also verified that the following DMC results are not strongly affected by the choice of procedure (see supplementary information for more details).

Table 3: Comparison of the mean unsigned error (MUE) in the dissociation energy obtained in this work using PPs, results obtained in this work using Gaussian09 and reoptimized orbitals (see text for details) with results from Ref. 7 using all-electron calculations with and without inclusion of scalar relativistic effects (see text for details).

method	this work using PPs	this work using PPs Gaussian09	Ref. 7	Ref. 7 scalar rel.
B3LYP	7.9	7.2	6.1	-
B97-1	6.3	5.9	5.1	4.9

A priori, this difference in MUE could be ascribed to a.) the pseudo-potential and b.) the basis set used since both can easily lead to errors of this order of magnitude as shown, for example, in Refs. 7 and 13. In an attempt to distinguish between basis set and PP errors, we applied the quadruple zeta basis set to five of the molecules with especially small experimental errors, namely CuCl, ZnH, ZnO, ZnS and ZnCl. For this basis set increase, the absolute energy of the molecules and transition metal atoms constituting these molecules dropped by 0.3 ± 0.2 kcal in the DFT. The average error in the dissociation energy, however, decreased only by 0.2 ± 0.1 kcal/mol. It is therefore likely that the increase in error between the results obtained here and in Ref. 7 are mainly due to our use of pseudo-potentials rather than due to the limited basis set we use. If the PPs are a source of error in DFT, this will also be true in DMC.

4.2 Optimization of the Jastrow function

The next step in obtaining DMC results is the optimization of the Jastrow function. As noted in Sec. 2.2 and motivated in Sec. 3, this optimization was done in a two step process, first minimizing the variance of the wavefunction and then using the resulting parameters in a linearized optimization of the energy. Using energy minimization, the ratio $\frac{E_{\mathrm{VMC}}-E_{\mathrm{UHF}}}{E_{\mathrm{DMC}}-E_{\mathrm{UHF}}}$ was above 85 % in all cases, showing that a large portion of the correlation energy captured in DMC is already present in variational Monte Carlo (VMC) after the optimization of the Jastrow function, which is in the range of typical values[17].

As we will demonstrate in the following our calculations clearly illustrate that energy minimization should be used for transition metal containing molecules even if the small core BFD-PPs are used, thus corroborating our findings from Sec. 3. First, we observe a clear difference in the trial wavefunctions obtained from energy and from variance optimization: for our database results we find that the energy optimization step decreases the VMC energy of the transition metals and transition metal compounds typically by more than 10 kcal/mol compared to results obtained after variance optimization. Increasing the quality of the variance optimization step by increasing the number of configurations used in the optimization to 200 000 , on the other hand, lowered the VMC energy only slightly. Energy optimization thus leads to considerably lower VMC energies. At the same time, the energy optimization increases the variance by more than 0.2 au^2. These two facts clearly indicate that the wavefunction obtained from energy minimization does not coincide with that found using variance minimization. Finding lower VMC energies for a certain trial wavefunction does not indicate by itself that the corresponding DMC calculations are also lower in energy. However, for the DMC(B3LYP) calculations, for which this was tested, the observed differences in energy prevail at the DMC level: For the atoms and molecules in the database, we found the DMC energies after energy minimization to be lower in energy by more than 2 kcal/mol, and often even by more than 6 kcal/mol. We thus conclude once more that the use of energy optimization is certainly favorable for accurate work with BFD pseudo-potentials.

4.3 DMC results

We now turn to the DMC energies themselves. To simplify the notation, DMC(x) will denote a DMC calculation for a trial wave function in which the Slater determinant is obtained with method x, where x denotes UHF, PBE, B3LYP and B97-1. First, we analyse the absolute energies and compare DMC energies for different trial wavefunctions. Figure 5 shows the DMC energies relative to the lowest DMC energy obtained.

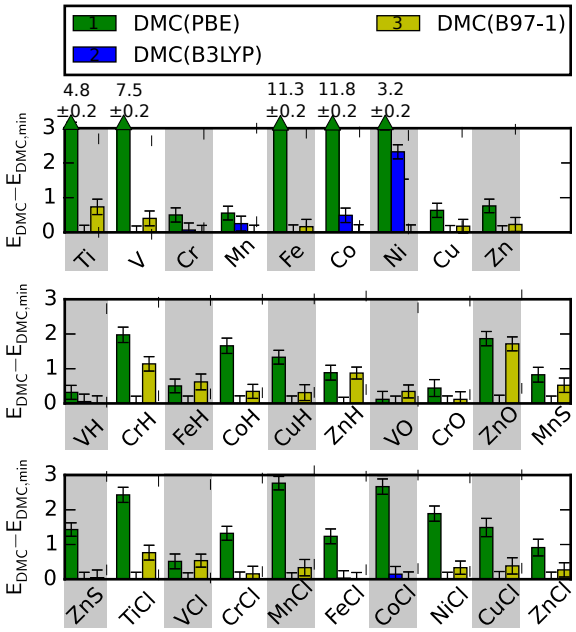

Figure 5: DMC absolute energies [kcal/mol] for different trial wavefunctions with respect to the lowest result obtained. All energies are given in kcal/mol. Errors that are larger than the scale chosen are given as numeric values. DMC(UHF) results are not shown, since they often do not fit the scale, especially for the molecues, where DMC(UHF) tends to give considerably higher total energies.

The unrestricted Hartree-Fock (UHF) trial wavefunction's performance is significantly worse than that of the DFT trial wavefunctions — a fact that has already been noted in Ref. 17 for transition metal oxides. Additionally, the performance of DMC(UHF) is significantly worse for the molecules than for the atoms, which is most likely due to an insufficient description of the d-p hybridization[15]. We note that in most cases where DMC(UHF) performs badly, the UHF calculation is quite strongly spin contaminated. The only exceptions to this rule (difference to best DMC result > 10 kcal/mol but no significant spin contamination in

the UHF result) are CrH and CrCl, which show a large error in DMC in spite of having a correct S^2 value in UHF.

For the DFT-trial wavefunctions studied (PBE, B3LYP and B97-1), the two hybrid functionals (B3LYP and B91-1) perform significantly better than the GGA functional PBE, following the general trend observed in the DFT results. However, except for V, Fe and Co, the differences between the QMC results obtained based on PBE, B3LYP and B97-1 with respect to the lowest energy DMC result obtained for a particular system are never larger than about 3 kcal/mol. For V, the initial PBE result is strongly spin contaminated ($S^2 = 4.20$ instead of $S^2 = 3.75$.). Such a strong spin contamination is not present in any of the other DFT calculations. Furthermore, assuming a ground state electronic configuration of $4s^2 3d^3$, the orbital which should nominally be the $4s$ orbital shows considerable mixing with the $3d_{zz}$ orbital — a fact that could be observed in several of the calculations. We also note that the observed difference DMC(PBE)-DMC(B3LYP) of approximately 7.5 kcal/mol is close to the first excitation energy ($^4F \rightarrow^6 D$) of V of approximately 6 kcal/mol[71]. In the case of Fe and Co, it is clear that the ground-state configuration found in PBE corresponds erroneously to $4s^1, 3d^7$ ($4s^1, 3d^8$) instead of $4s^2, 3d^6$ ($4s^2, 3d^7$), which is correctly found in B3LYP and B97-1 (actually, a similar observation applies to Ni). Extensive tests reordering the guess orbitals and using different convergence schemes available in the GAMESS-US code did not yield a better energy. It is likely that the first excited state of Co ($4s^1, 3d^8$; a^4F) is found in PBE instead of the ground state ($4s^2, 3d^7$; b^4F) (the first excitation energies tabulated in the NIST database[71] lie between 10 kcal/mol and 14 kcal/mol depending on J; DMC(PBE)-DMC(B3LYP) $\approx$ 11.8 kcal/mol). For Fe, however, the difference DMC(PBE)-DMC(B3LYP) $\approx$ 11.5 kcal/mol is considerably smaller than the first excitation energy ($^5D \rightarrow^5 F$), given by nearly 20 kcal/mol[71]. It should be kept in mind that these "outliers" will influence the following analysis.

DMC(B3LYP) and DMC(B97-1), the two DMC calculations based on trial wavefunctions obtained with hybrid functionals, perform similarly well. In terms of absolute DMC

energy, they give the lowest energy for all atoms and molecules and thus clearly outperform DMC(UHF) and, to a lesser extent, DMC(PBE). The differences in absolute DMC energy for trial wavefunctions based on the two different hybrid functionals tested are generally small. The good performance of the DMC based on hybrid functionals for the database of TMCDs does not come unexpectedly as hybrid functionals are known to outperform GGA functionals at the DFT level for molecular systems.

Judging by the absolute energies, we expect DMC with a B3LYP or B97-1 Slater-Jastrow trial wavefunction to perform best also with respect to the dissociation energies. This expectation is indeed met as can be seen in Fig. 4 (note also the comparison with the DFT results in the same figure). The corresponding mean unsigned and signed errors (MUE and MSE) are given in Tab. 4 (and Fig. 6) together with the corresponding standard error of the mean unsigned error

$$\sigma(\mathrm{MUE}) = \sqrt{\sum_{i=1}^{N} \frac{\left(|D_e - D_e^{\mathrm{exp.}}| - \mathrm{MUE}\right)^2}{N \cdot (N-1)}}, \tag{10}$$

where $N = 20$. The additional uncertainties in the values of the MUE and the MSE that arise from the uncertainties in the experimental results and the stochastic nature of QMC are given by the value-range in the MUE and MSE column of Tab. 4. These uncertainties, which scale as $\mathcal{O}(1/\sqrt{N})$ in the MUE and MSE, are dominated by the experimental uncertainties, since the stochastic uncertainty in the DMC calculation of each molecule was converged to significantly less than half a kcal/mol, which is considerably lower than most experimental uncertainties.

As expected, DMC(B3LYP) and DMC(B97-1) perform best for the dissociation energies, with a MUE of 5.3 kcal/mol and 5.4 kcal/mol respectively. Taking the standard deviation of the mean absolute error into account (see Table 4), it is basically impossible to say whether DMC(B97-1) or DMC(B3LYP) performs better. However, even for these best performing trial wavefunctions, MUEs obtained for this database of TMCDs are considerably larger than

Table 4: Mean unsigned and signed error (MUE, MSE) in the dissociation energy (Eq. (9)) and the standard deviation of the mean unsigned error σ(MUE) evaluated for the set of 20 transition metal containing molecules given in Tab. 2 (see text for details). The last result for DMC(B3LYP) with the note "Padé func." was obtained using a Jastrow factor parametrized by Padé functions in QWALK, while all other DMC calculations were obtained with polynomial Jastrow functions in CASINO. Negative values in the MSE indicate underbinding.

method	MUE [kcal/mol]	σ(MUE) [kcal/mol]	MSE [kcal/mol]
UHF	35.2 ± 0.3	7.0	-35.2 ± 0.3
PBE	8.4 ± 0.3	1.6	7.4 ± 0.3
B3LYP	7.9 ± 0.3	1.0	-3.2 ± 0.3
B97-1	6.3 ± 0.3	1.0	-2.7 ± 0.3
DMC(UHF)	11.2 ± 0.4	2.8	-10.0 ± 0.4
DMC(PBE)	6.4 ± 0.4	1.0	0.5 ± 0.4
DMC(B3LYP)	5.3 ± 0.4	0.6	-2.3 ± 0.4
DMC(B97-1)	5.4 ± 0.4	0.7	-2.7 ± 0.4
DMC(B3LYP) Padé func.	4.5 ± 0.4	0.6	-1.7 ± 0.4

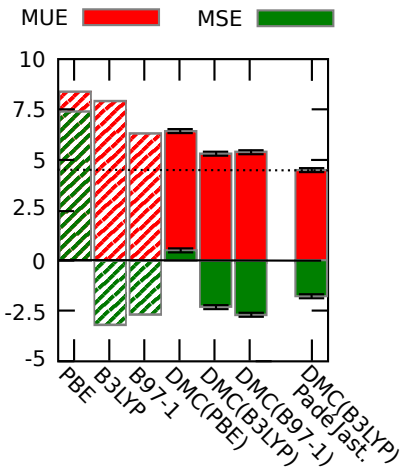

Figure 6: Graphical representation of mean unsigned and signed errors (MUE, MSE) given in Tab. 4 for wavefunctions based on DFT calculations. DFT results are hatched. The dotted line shows the location of the best DMC result obtained (using Padé style Jastrow functions — see Sec. 4.3.2).

those obtained in DMC for the G2 set: Using a triple zeta basis, a single Slater determinant trial wavefunction with Hartree-Fock orbitals, and all-electron potentials, the mean unsigned error for the G2 set is given by MUE = 3.2 kcal/mol[65]. This demonstrates the difficulty associated with the accurate determination of dissociation energies for TMCDs.

Since the database is limited to only 20 molecules and the MUE are thus subject to an uncertainty quantified by σ(MUE) (Eq. 10), we use statistical hypothesis testing to calculate the probability p that DMC performs better than the best performing DFT functional used as

$$p = \int_{-\infty}^{Z} N(x)\mathrm{d}x, \tag{11}$$

where $N(x)$ denotes a normal distribution and

$$Z = \frac{\mu_{\mathrm{DFT}} - \mu_{\mathrm{DMC}}}{\sqrt{\sigma_{\mathrm{DFT}}^2 + \sigma_{\mathrm{DMC}}^2}}. \tag{12}$$

The values for μ and σ are taken from the MUE and σ(MUE) values in Tab. 4. Carrying out this analysis, we obtain a probability of about 80 % for DMC(B3LYP) to perform as well or better than the best performing DFT result, DFT(B97-1) (70 %, if the reoptimized DFT results stated in Tab. 3 are used). The corrected sample variance v of the unsigned error, given by $v = N \cdot \sigma(\mathrm{MUE})^2$ (see Tab. 4), is lower for DMC than for DFT, indicating the presence of fewer "outliers" in DMC, i.e. systems with unexpectedly higher errors compared to the average error obtained.

For DMC(PBE), the MUE is significantly higher than that of DMC(B97-1) and DMC(B3LYP) and even lies slightly above that of the DFT(B97-1) results. This clearly shows that, if the DMC calculations are based on GGA functionals (which do not perform particularly well for this database), the DFT calculations that are based on hybrid functionals can compete with DMC for this database of TMCDs. A similar conclusion was drawn in Ref. 7 for DFT compared to CCSD(T), where it was shown that DFT can even outperform CCSD(T) for this particular database. However, in spite of this good performance of hy-

brid DFT, it should be noted that the errors of DFT(PBE) are significantly higher than those of DMC(PBE). Furthermore, for DMC, the MUE increases only by 1.0 kcal/mol from DMC(B97-1) to DMC(PBE), while for DFT the MUE increases by 2.1 kcal/mol for the same functional change. This shows that the errors in DMC are not so strongly dependent on the XC-functional used to generate the trial wavefunction. The weak dependence on the XC-functional becomes even more evident when the molecules containing V, Fe or Co, which were marked as "outliers" in PBE, are excluded from the analysis. Then, the MUE of the remaining 13 molecules for DMC(PBE) is reduced to 5.1 kcal/mol, while those of DMC(B97-1) and DMC(B3LYP) decrease only slightly to 5.3 kcal/mol and 5.1 kcal/mol respectively.

The fact that mean errors in the DMC results depend only weakly on the functional used for the trial wavefunction generation is a very important result: As also obvious from the present results, hybrid functionals often allow a comparatively good description of molecular properties. Their use is, however, problematic in metallic systems[72–75]. In systems, such as heterogeneous catalysis, where both, the molecular interaction and the metallic surface are important, this limits the performance of hybrid DFT in spite of its good performance for molecules. DMC calculations, which seem to depend less on the XC-functional used in the trial function, may thus well amend some of the problems hybrid DFT encounters in such systems (see Refs. 72–75, and also Refs. 76,77, which mention these problems). Differently phrased, since the performance of DMC(PBE) is as good as that of hybrid functionals for the TMCDs, the use of DMC(PBE) for molecule-metal surface systems might lead to a similar increase of accuracy as that observed when changing from the use of standard GGA functionals like PBE to that of hybrid functionals in the description of molecular systems.

Looking at the errors obtained for each molecular species in more detail (see Fig. 4a and Fig. 4b), an interesting feature of the DMC results strikes the eye: putting the previously discussed outliers aside, the errors in the DMC calculations seem to be nearly constant for different functionals used in the trial wavefunction generation for most of the molecules. This is in stark contrast to the DFT results, where the errors often even have different

signs depending on the functional, and where even the two hybrid functionals often give astonishingly different results. To put this observation onto more solid ground, we calculate the mean of the corrected sample variance of the dissociation energy of the molecules in the data set for different XC-functionals

$$v^*_{\mathrm{DMC}} = \frac{\sum_{i=1}^{N_{\mathrm{mol}}} \sum_{j=1}^{N_{\mathrm{xc}}} \left(D_e^{\mathrm{DMC}}(i;j) - \bar{D_e}^{\mathrm{DMC}}(i) \right)^2}{N_{\mathrm{mol}} N_{xc} - N_{\mathrm{mol}}}, \tag{13}$$

where i runs over the $N_{\mathrm{mol}} = 20$ molecules, j runs over the $N_{\mathrm{xc}} = 3$ XC-functionals and $D_e^{\mathrm{DMC}}(i,j)$ denotes the dissociation energy of molecule i calculated with DMC using a trial wavefunction generated with one of the XC-potentials $j \in$ [PBE, B3LYP, B97-1]. The quantity $\bar{D_e}^{\mathrm{DMC}}(i)$ is the average of $D_e^{\mathrm{DMC}}(i,j)$ over all j. The corresponding value for DFT is calculated by replacing the dissociation energies D_e^{DMC} with the results from DFT. For DFT we obtain $v^*_{\mathrm{DFT}} = 65\,\mathrm{kcal}^2/\mathrm{mol}^2$, whereas $v^*_{\mathrm{DMC}} = 9\,\mathrm{kcal}^2/\mathrm{mol}^2$. Disregarding the results including V, Fe and Co, for which the PBE results were marked as outliers as discussed above, the values change to $v^*_{\mathrm{DFT}} = 46\,\mathrm{kcal}^2/\mathrm{mol}^2$ and $v^*_{\mathrm{DMC}} = 0.5\,\mathrm{kcal}^2/\mathrm{mol}^2$ respectively. These numbers show clearly that not only does the average error of DMC depend less on the XC-functional used in the trial wavefunction generation, but also the actual DMC errors are more or less independent of the functional, which is in strong contrast to the DFT results .

The remaining discrepancies between DMC and experiment can have several different origins: 1.) the experimental values come with error bars of up to 2 kcal/mol, 2.) although the locality error was shown to cancel out to a high degree in CuH if BFD PPs are used, the influence of the locality error may still be present in other molecules 3.) the fixed-node approximation may lead to imprecise results (including errors due to a restricted basis in the trial wavefunction generation) and 4.) the pseudo-potentials may not exhibit the required transferability in such systems.

We will investigate each of these influences in more detail in the following.

4.3.1 Influence of uncertainties in the experimental results

Since the uncertainty in the experimental results is on the same scale as the MUE, these errors may considerably influence the results, especially since the DMC results are so strongly consistent for different trial wavefunction generations. To check the influence of the experimental error, we analyzed a smaller database of 9 molecules with an experimental error smaller than 1.2 kcal/mol. This value was chosen such that the experimental error is considerably smaller than that of the total database, but that it retains the ratio of molecules with strong single- and multi-determinant character compared to the full set. In most cases the MUE did not change significantly. We thus conclude that the experimental quality of the database is sufficient — a conclusion that was also drawn in Ref. 7, based on different arguments.

4.3.2 Influence of residual locality errors

To test for the influence of the locality approximation on the above results, we switch from the CASINO code to QWALK. This allows us to change the computational approaches (most importantly, the form of the Jastrow factor and the optimization techniques; for details see Sec. 2.2) much more drastically than just going to a larger parametrization of the Jastrow function, and permits us additionally to test for code related problems. Reanalyzing the DMC(B3LYP) results with QWALK using its standard Jastrow factor given by Padé functions, we find an improved MUE of 4.5 kcal/mol, compared to the MUE of 5.3 kcal/mol, which we found using a polynomial form in the Jastrow function in CASINO (see Tab. 4). Since both calculations use the same PPs and the same Slater function in the trial wavefunctions, this improvement can only be due to the changes in the Jastrow functions, and, therefore, the locality approximation. Obviously, the mean error in the dissociation energies is not only determined by the fixed-node and PP errors, but also by locality errors.

To allow a more detailed investigation, we show in Fig. 7 the differences obtained in absolute energies when using a polynomial form of the Jastrow function (CASINO) and Padé

functions (QWALK). Especially for the heavier transition metal elements and their com-

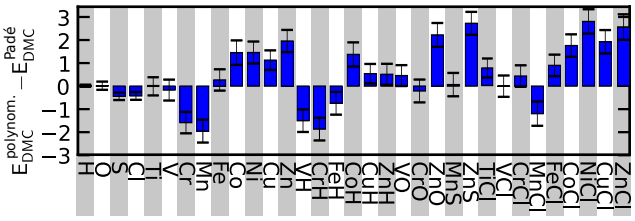

Figure 7: Difference in absolute DMC energies [kcal/mol] obtained using a polynomial Jastrow (CASINO) and Padé functions (QWALK). Positive values correspond to systems where the QWALK result is lower in energy.

pounds, the Padé function parametrization tends to give a lower DMC result, suggesting that for these elements the structure of the Jastrow factor used in QWALK is more appropriate. For ZnO, ZnS, ZnCl and NiCl, the differences in absolute energy between the two parametric forms are even greater than 2 kcal/mol. We can therefore conclude that, at least for some TMCDs, the locality approximation adds a systematic uncertainty to the results which is larger than the error bars aspired to when seeking chemical accuracy.

In the end, we are, however, only interested in relative energies and the change in locality error may be consistent among the atoms and molecules and thus cancel out. In Figs. 8 and 9, we therefore investigate the dissociation energy itself. Fig. 8 shows the difference in dissociation energy between the CASINO results with polynomial Jastrow form and the QWALK results, which use Padé functions. This gives an idea on how stable the results are when using different forms of the Jastrow function. In Fig. 9, the obtained dissociation energies are compared to the experimental values, allowing a judgment of the overall quality of the results. When analyzing this graph, one should keep in mind that for any particular molecule, the interplay of the locality, fixed-node and PP related errors can lead to error cancellation in the dissociation energy between different types of error — an error cancellation that is actually not desirable since it may not be grounded on physical similarities between the atomic and the molecular system, but on pure chance. A reduction in the error of a specific dissociation energy does therefore not necessarily go along with an improvement of the corresponding Jastrow functions. Making this correlation is only possible for total

energies.

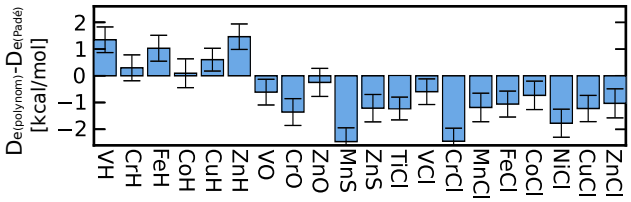

Figure 8: Difference in dissociation energies when using a polynomial Jastrow (CASINO) or Padé functions (QWALK). (Energies in kcal/mol).

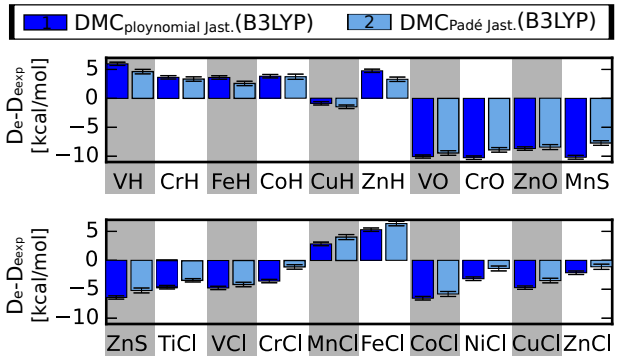

Figure 9: Comparison of dissociation energies calculated for a polynomial Jastrow (CASINO) and Padé functions (QWALK) with the experimental results. (Energies in kcal/mol). Negative values denote underbinding, positive ones, overbinding.

Analyzing Fig. 8, we note first of all that for some molecules like ZnO, ZnS, ZnCl and NiCl, for which the total energies differ strongly, a large part of the Jastrow-related energy change cancels out when calculating energy differences. On the other hand, we also observe that the opposite is true for other molecules and that large differences in the dissociation energy of two or more kcal/mol are observed for some of the TMCDs, namely for TiCl, CrCl and MnS. We can thus conclude that, for the dissociation energies as well as the total energies of TMCDs described by DMC using single-Slater-Jastrow trial wavefunctions, we have to deal with systematic uncertainties related to the Jastrow parametrization that are in the order of considerably more than 1 kcal/mol. At this point, it should be stressed once more that the presence of comparatively large locality errors is not a flaw of DMC in itself: first of all, many systems, especially those only containing first row elements, will suffer from considerably smaller locality errors and second, the trial wavefunction can be

improved, thereby reducing the fixed-node and locality error. The use of multi-determinant trial wavefunctions and orbital optimization is, while not entirely straightforward due to the necessity of good optimization schemes, common practice for small systems. For the large systems, which we ultimately have in mind, this is, however, not feasible. The awareness of the possible influence of the locality error on the accuracy of the results and an estimate of its size are thus important.

Although the locality errors we observe are large compared to chemical accuracy, Fig. 9 clearly suggests that it is nonetheless not the locality error which dominates the errors in the dissociation energies: the change of Jastrow parametrization only seems to add a comparatively small fluctuation in the error to an overall often large error bar. In order to decrease this uncertainty, it might be desirable to eliminate or decrease the uncertainties coming from the locality approximation as much as possible and thus to optimize the Jastrow factor as well as possible. We therefore re-investigated the cases for which the largest dependence on the Jastrow function was found, namely TiCl, CrCl and MnS, with larger Jastrow parametrizations: in CASINO we go from order 4 in the polynomial describing u and χ to order 6 and in the f-term we go to an order 3x3x3 in the three variables. In QWALK we go from an f-function with 12 terms to an f-function with 30 terms. The corresponding results are given in detail in the supplementary information. For the three systems considered, the added flexibility obtained through the larger Jastrow expansions influenced the result on the order of about 1 kcal/mol, but it did not systematically lead to smaller deviations between the computed and the experimental dissociation energies, for both parametric forms.

We therefore conclude our analysis on the Jastrow parametrization by emphasizing that our results, while not yet allowing a final elimination of the locality error within the single-Slater Jastrow framework, clearly show that the use of single-Slater-Jastrow trial wavefunctions in DMC can and will lead to considerable uncertainties (of the order of several kcal/mol) in the dissociation energies of certain TMCDs due to the presence of locality errors. On the other hand, the large deviations observed between the calculated and the experimental re-

sults (e.g. for MnS) seem to be dominated by errors other than the locality error (fixed-node or PP related errors), since the deviations from the experimental values observed in the dissociation energy are considerably larger than the fluctuations observed when varying the Jastrow function.

4.3.3 Influence of fixed-node errors and PP related errors

We now turn to the fixed-node and the pseudo-potential errors. The fixed-node error may be considered as controllable to a certain extent in small systems, since the trial wavefunction can be improved by adding more determinants. An improvement of the trial wavefunction will also decrease the locality errors discussed in the previous section. In large systems, however, the fixed-node error becomes uncontrollable, since there is so far no way to improve the trial wavefunction systematically until convergence is reached, although embedding schemes may provide a step in this direction[78].

The fixed-node error might be expected to be smaller for molecules with a strong single-determinant character[7]. Following Ref. 7, we split the database into 7 molecules with strong single-determinant character (CuCl, ZnH, ZnS, ZnCl, CrCl, MnCl, FeCl) and 13 molecules with multi-determinant character. The resulting MUE is shown in Table 5. For molecules with single-determinant character wavefunctions, the DMC errors are clearly smaller, suggesting that a portion of the remaining error is indeed due to the fixed node approximation. This is especially true for DMC(UHF), where the MUE decreases from 14.1 kcal/mol for molecules with multi-determinant character to 5.8 kcal/mol for molecules with single determinant character — a value close to the DMC(PBE) result. On the other hand, the decrease in error is considerably smaller in DMC than in DFT. This may indicate that the pseudo-potentials do really constitute another source of error, which would also explain the consistency in the DMC results for different trial wavefunctions: a PP error would introduce a systematic error that is different for each molecular system, but might not much depend on the exchange-correlation functional used in obtaining the trial wavefunction. In terms

Table 5: Mean unsigned error (MUE) in the dissociation energy, when separating the molecules in the database in to those with strong single- and those with strong multi-determinant character of the wavefunction.

method	sing. ref.	mult. ref.	all
	MUE [kcal/mol]	MUE [kcal/mol]	MUE [kcal/mol]
UHF	17.7	44.6	35.2
PBE	3.5	11.0	8.4
B3LYP	6.0	9.0	7.9
B97-1	4.1	7.4	6.3
DMC(UHF)	5.8	14.1	11.2
DMC(PBE)	5.5	6.9	6.4
DMC(B3LYP)	4.2	5.9	5.3
DMC(B97-1)	4.2	6.0	5.4
DMC(B3LYP) Padé func.	3.5	5.0	4.5
# of molecules	7	13	20

of size of error, PP errors which are of the same order of magnitude of the errors obtained, have been observed in earlier DMC calculations (e.g. Refs. 13,14,16,67,79). One way to improve on these errors would be to use He-core PPs[79]. Another possibility would be to try to improve standard DFT-PPs by adding more angular momentum channels[80] or to improve the transferability of the PPs to different chemical environments by designing the PPs in a different way (e.g. 81,82).

Although PP errors may contribute to the observed errors, fixed-node errors are sure to come into play on the accuracy scale we aim to reach. For accuracies at and around chemical accuracy, even small weight excitations can be extremely relevant (see for example Refs. 83–87). In order to distinguish further between fixed-node and pseudo-potential errors, we therefore performed multi-determinant DMC calculations and DMC calculations including backflow transformations[59] for Cu, CuH and MnS and its constituents. We chose these molecules, since Cu and CuH were already analyzed in detail in Sec. 3. MnS was chosen as an example of a molecule showing a large deviation from the experimental results in DMC.

All calculations were performed at a time-step $\tau = 0.002\,\mathrm{au}$, without zero-timestep ex-

trapolations. However, the remaining corrections should be small.

We start our discussion with Cu and CuH. The results from multi-determinant expansions and when using backflow are shown in Fig. 10.

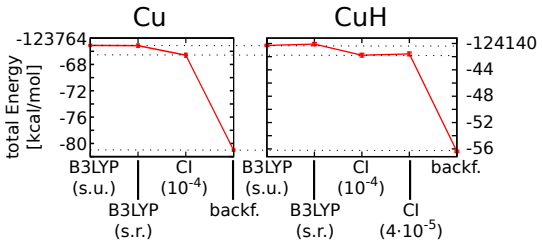

Figure 10: Absolute DMC energies obtained for Cu and CuH using spin unrestricted (s.u.) and spin restricted (s.r.) B3LYP Slater-Jastrow trial wavefunctions, a CI expansion up to the minimal weight specified, or backflow transformations on the unrestricted B3LYP trial wavefunction.

For Cu, we used all determinants with a weight larger than 1×10^{-4}. This corresponds to a sum weight of 94.8 %. In DMC, this lead to an energy decrease of about 1.4 kcal/mol. A very similar decrease was observed for the CI expansion of CuH using 47 (weight $> 1 \times 10^{-4}$, sum weight 93.4 %) and 201 (weight $> 4 \times 10^{-5}$, sum weight 94.2 %) determinants. The backflow transformation lead to a much larger decrease in DMC energy of about 16 kcal/mol, much of which is most likely related to the core. The energy decrease was again very similar for Cu and CuH, indicating that the fixed node error, while being quite large, cancels out quite well.

Tests using backflow for MnS also showed the absolute energies to decrease by several kcal/mol, but no relevant change in the dissociation energy was observed. This observation is very similar to that made for CuH. Further test for compounds showing a large discrepancy with the experimental values would be desirable, in order to distinguish further between fixed-node and pseudo-potential errors.

5 Summary and Conclusion

We present a systematic investigation of the performance of diffusion Monte Carlo (DMC) for transition metals and transition metal containing dimers (TMCDs). Studying the database of twenty 3d-TMCDs, we were able to gauge the accuracy of DMC for such systems. With a mean unsigned error (MUE) of 5.3 kcal/mol (4.5 kcal/mol when using the structure of Jastrow which is default in QWALK) our DMC error lies below the MUE we obtain with DFT with pseudo-potentials. The best performing XC-functional tested (B97-1) shows a MUE which is 1 kcal/mol higher than that of the best performing DMC calculation, DMC(B3LYP). The accuracy of our DMC results is close to that of all-electron, scalar relativistic CCSD(T) and the best all-electron DFT results obtained in Ref. 7 (4.5 kcal/mol — in the case of CCSD(T) using a very extensive basis set).

Our DMC results exhibit several favorable characteristics: first, the MUE of the DMC results has a smaller sample variance than the MUE of DFT. This indicates the presence of fewer cases with exceptionally high (or low) errors. DMC thus has fewer unpredictable "outliers" that make it difficult to judge the quality of a specific result. Furthermore, the errors we obtain for each individual molecule are very robust against changes in the XC-functional used to generate the trial wavefunction. This observation, and the fact that the use of trial functions based on GGA functionals does not strongly degrade the quality of the DMC results, are extremely relevant for molecule-metal surface systems relevant to heterogeneous catalysis. For these systems one does not only need to describe the molecule to very high precision but also the metallic surface. This is difficult to achieve with hybrid functionals which generally perform very well for molecular systems but suffer from difficulties for metals and furthermore exhibit an increased computational cost in DFT calculations.

In the context of this work, we performed an extensive error analysis. Our work highlights the impact of the use of pseudo-potentials on the results and the possible failure of the locality approximation, a source of error that is often overlooked. Investigations focusing on Cu and CuH show that the locality approximation can lead to non-systematic errors of several (tens

of) kcal/mol in both absolute and relative energies if no care is taken in the construction of the Jastrow factor, which is used in the evaluation of the non-local part of the PP. Our data allow us to draw conclusions regarding best practices for treating systems which are prone to locality errors, such as the TMCDs considered here: first of all, we show that the use of the Ne core pseudo-potentials (PPs) by Burkatzki, Filippi and Dolg[9,10] can greatly reduce the locality error in Cu and CuH compared to the other PPs tested. Furthermore, the use of large Jastrow parametrizations is strongly favorable, especially if large core PPs are used. The use of T-moves instead of the locality approximation introduced in Refs. 23,24 helps reducing errors in DMC arising from the approximate treatment of the non-local potential. Regarding the optimization of the Jastrow function, we give evidence that trial wavefunctions optimized using linear energy minimization give considerably better DMC results for the dissociation energy of CuH than trial wavefunctions optimized by unreweighted variance minimization, especially if the T-move scheme is not used. Further evidence for the superiority of trial wavefunctions which are optimized with respect to energy for TMCDs, is given by the fact that lower DMC energies are obtained for the entire database of 20 molecules for the trial wavefunctions optimized by minimizing the energy.

Regarding the achievable accuracy of DMC for TMCDs, our results indicate that locality errors, remaining in spite of the effort made to keep them small, as well "physical" shortcomings of the PPs due to limits in their transferability and the missing core-valence interaction, add to the fixed-node error and hence to the mean error observed for the TMCDs. The use of He core PPs or PPs with improved transferability may therefore further reduce the MUE obtained in DMC. In order to address the remaining discrepancies between experimental data and DMC results and in order to reduce the observed MUE, we encourage further investigations aimed at separating the fixed-node error from pseudo-potential errors.

Taking the performance of a method in describing small molecules as an indicator of its ability to correctly describe interactions in solids and at transition metal surfaces, we expect the present work to establish DMC further as an interesting candidate for calculations on

large or periodic systems in which the interaction with transition metals is important. For such systems, higher level quantum chemistry methods cannot be applied directly any more due to the large size of the system. Quantum chemistry calculations of systems embedded in a DFT environment[78] might still be feasible, but such methods come with additional challenges due to the requirement to perform the embedding correctly and accurately. DFT, on the other hand, while being scalable to large systems, often does not yield accurate enough results and depends strongly on the XC-functional. Since we found the error in DMC for TMCDs to depend only weakly on the functional used to generate the trial wavefunction, DMC may be of interest especially for systems such as transition metal surfaces interacting with molecules, where the hybrid DFT functionals that perform best for the molecular database offer no improvement over GGA functionals.

Our results provide useful guidelines for achieving greater accuracy in DMC calculations on transition metal containing systems, such as the use of small core BFD-PPs[b], of large Jastrow functions, and of energy optimization of the trial wavefunction, and provide insights in the accuracy achievable with QMC. We expect these results to be valuable in view of the increasing interest in treating strongly correlated, transition metal containing systems in DMC (e.g. Ref. 15,19,20,88), and the inherent relevance of transition metals to catalysis and superconducting materials.

Acknowledgement

The authors thank Neil Drummond for helpful discussions.

This work was sponsored by NWO Exacte Wetenschappen, EW (NWO Physical Sciences Division) for the use of supercomputer facilities and with financial support from the European Research Council through an ERC-2013 Advanced Grant No. 338580. L.K.W. was supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific

[b]We note that other small core PPs, such as the Stuttgart-Köln PPs have not been tested in the course of this work, but may in principle also be applied in diffusion Monte Carlo calculations.

Computing Research, Scientific Discovery through Advanced Computing (SciDAC) program under Award Number FG02-12ER46875.

Supporting Information Available

A.) Full data and discussion for DMC results comparing different parametric forms and sizes for TiCl, CrCl and MnS. B.) DFT and DMC results for a second strategy to obtain the DFT wavefunctions. This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Pierloot, K. *Int. J. Quantum Chem.* **2011**, *111*, 3291–3301.

(2) Pou-Amérigo, R.; Merchán, M.; Nebot-Gil, I.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1994**, *101*, 4893–4902.

(3) Blomberg, M. R. A.; Siegbahn, P. E. M.; Roos, B. O. *Mol. Phys.* **1982**, *47*, 127–143.

(4) Goel, S.; Masunov, A. E. *Int. J. Quantum Chem.* **2011**, *111*, 4276–4287.

(5) Kulik, H. J.; Marzari, N. In *Fuel Cell Science: Theory, Fundamentals and Biocatalysis*; Wieckowski, A., Norskov, J. K., Eds.; John Wiley & Sons, Inc., 2010; pp 433–455.

(6) Jensen, K. P.; Roos, B. O.; Ryde, U. *J. Chem. Phys.* **2007**, *126*, 014103.

(7) Xu, X.; Zhang, W.; Tang, M.; Truhlar, D. G. *J. Chem. Theory Comput.* **2015**, *11*, 2036–2052.

(8) Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2005**, *123*, 064107.

(9) Burkatzki, M.; Filippi, C.; Dolg, M. *J. Chem. Phys.* **2007**, *126*, 234105.

(10) Burkatzki, M.; Filippi, C.; Dolg, M. *J. Chem. Phys.* **2008**, *129*, 164115.

(11) Scemama, A.; Applencourt, T.; Giner, E.; Caffarel, M. *J. Chem. Phys.* **2014**, *141*, 244110.

(12) Thomas, R. E.; Booth, G. H.; Alavi, A. *Phys. Rev. Lett.* **2015**, *114*, 033001.

(13) Petz, R.; Lüchow, A. *ChemPhysChem* **2011**, *12*, 2031–2034.

(14) Bande, A.; Luchow, A. *Phys Chem Chem Phys* **2008**, *10*, 3371–3376.

(15) Wagner, L. K. *J. Phys. Condens. Matter* **2007**, *19*, 343201.

(16) Wagner, L. K.; Mitas, L. *J. Chem. Phys.* **2007**, *126*, 034105.

(17) Wagner, L.; Mitas, L. *Chem. Phys. Lett.* **2003**, *370*, 412 – 417.

(18) Ovcharenko, I. V.; Jr, W. A. L.; Xiao, C.; Hagelberg, F. *J. Chem. Phys.* **2001**, *114*, 9028–9032.

(19) Foyevtsova, K.; Krogel, J. T.; Kim, J.; Kent, P.; Dagotto, E.; Reboredo, F. A. *Phys. Rev. X* **2014**, *4*, 031003.

(20) Wagner, L. K.; Abbamonte, P. *Phys. Rev. B* **2014**, *90*, 125129.

(21) Williamson, A. J.; Hood, R. Q.; Grossman, J. C. *Phys. Rev. Lett.* **2001**, *87*, 246406.

(22) Alfè, D.; Gillan, M. J. *J. Phys.: Condens. Matter* **2004**, *16*, L305.

(23) Hurley, M. M.; Christiansen, P. A. *J. Chem. Phys.* **1987**, *86*, 1069.

(24) Mitáš, L.; Shirley, E. L.; Ceperley, D. M. *J. Chem. Phys.* **1991**, *95*, 3467.

(25) Lester Jr., W. A.; Mitas, L.; Hammond, B. *Chemical Physics Letters* **2009**, *478*, 1–10.

(26) Foulkes, W. M. C.; Mitas, L.; Needs, R. J.; Rajagopal, G. *Rev. Mod. Phys.* **2001**, *73*, 33–83.

(27) Kolorenč, J.; Mitas, L. *Rep. Prog. Phys.* **2011**, *74*, 026502.

(28) Toulouse, J.; Umrigar, C. J. *J. Chem. Phys.* **2007**, *126*, 084102.

(29) Lüchow, A. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 388–402.

(30) Ceperley, D. *J. Stat. Phys.* **1986**, *43*, 815–826.

(31) Hammond, B. L.; Reynolds, P. J.; Lester, W. A. *J. Chem. Phys.* **1987**, *87*, 1130–1136.

(32) ten Haaf, D. F. B.; van Bemmel, H. J. M.; van Leeuwen, J. M. J.; van Saarloos, W.; Ceperley, D. M. *Phys. Rev. B* **1995**, *51*, 13039–13045.

(33) Casula, M. *Phys. Rev. B* **2006**, *74*, 161102.

(34) Casula, M.; Moroni, S.; Sorella, S.; Filippi, C. *J. Chem. Phys.* **2010**, *132*, 154113.

(35) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993–2006.

(36) GGA_FHI — Abinit, (accessed: Nov. 11, 2014). `http://www.abinit.org/downloads/psp-links/gga_fhi`, 2015.

(37) Fuchs, M.; Scheffler, M. *Computer Physics Communications* **1999**, *119*, 67–98.

(38) Trail, J. R.; Needs, R. J. *J. Chem. Phys.* **2005**, *122*, 174109.

(39) Pseudopotential and basis set library (accessed: Dec. 9, 2014). `https://vallico.net/casinoqmc/pplib/`, 2015.

(40) Dolg, M.; Filippi, C. Private communication. 2015.

(41) Krauss, M.; Stevens, W. J. *Annu. Rev. Phys. Chem.* **1984**, *35*, 357–385.

(42) Pacios, L. F.; Calzada, P. G. *Int. J. Quantum Chem.* **1988**, *34*, 267–277.

(43) Stevens, W. J.; Krauss, M.; Basch, H.; Jasien, P. G. *Can. J. Chem.* **1992**, *70*, 612–630.

(44) Mitáš, L. *Phys. Rev. A* **1994**, *49*, 4411–4414.

(45) Lee, Y.; Kent, P. R. C.; Towler, M. D.; Needs, R. J.; Rajagopal, G. *Phys. Rev. B* **2000**, *62*, 13347–13355.

(46) Cao, X.; Dolg, M. *WIREs Comput Mol Sci* **2011**, *1*, 200–210.

(47) Dolg, M.; Cao, X. *Chem. Rev.* **2012**, *112*, 403–480.

(48) Stuttgart/Köln table of PPs (accessed: Mar 22, 2016). `http://www.tc.uni-koeln.de/PP/index.en.html`.

(49) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

(50) DeFusco, A. Private communication: code maintained by Albert DeFusco. 2015.

(51) Needs, R. J.; Towler, M. D.; Drummond, N. D.; Ríos, P. L. *J. Phys.: Condens. Matter* **2010**, *22*, 023201.

(52) Giannozzi, P. et al. *J. Phys. Condens. Matter* **2009**, *21*, 395502.

(53) Wagner, L. K.; Bajdich, M.; Mitas, L. *Journal of Computational Physics* **2009**, *228*, 3390–3404.

(54) Needs, R. J.; Towler, M.; Drummond, N.; Riós, P. L. *Casino User's Guide Version 2.13*; 2014.

(55) Wagner, L. K. Qwalk manual (accessed Feb. 2, 2016). `http://www.qwalk.org/docs/`, 2016.

(56) Umrigar, C. J.; Wilson, K. G.; Wilkins, J. W. *Phys. Rev. Lett.* **1988**, *60*, 1719–1722.

(57) Drummond, N. D.; Needs, R. J. *Phys. Rev. B* **2005**, *72*, 085124.

(58) Umrigar, C. J.; Toulouse, J.; Filippi, C.; Sorella, S.; Hennig, R. G. *Phys. Rev. Lett.* **2007**, *98*, 110201.

(59) López Rios, P.; Ma, A.; Drummond, N. D.; Towler, M. D.; Needs, R. J. *Phys. Rev. E* **2006**, *74*, 066701.

(60) Umrigar, C. J.; Nightingale, M. P.; Runge, K. J. *J. Chem. Phys.* **1993**, *99*, 2865–2890.

(61) Flyvbjerg, H.; Petersen, H. G. *J. Chem. Phys.* **1989**, *91*, 461.

(62) Lee, R. M.; Conduit, G. J.; Nemec, N.; López Ríos, P.; Drummond, N. D. *Phys. Rev. E* **2011**, *83*, 066706.

(63) Wang, F.; Ziegler, T.; van Lenthe, E.; van Gisbergen, S.; Baerends, E. J. *J. Chem. Phys.* **2005**, *122*, 204103.

(64) Nemec, N.; Towler, M. D.; Needs, R. J. *J. Chem. Phys.* **2010**, *132*, 034111.

(65) Petruzielo, F. R.; Toulouse, J.; Umrigar, C. J. *J. Chem. Phys.* **2012**, *136*, 124116.

(66) Esler, K. P.; Cohen, R. E.; Militzer, B.; Kim, J.; Needs, R. J.; Towler, M. D. *Phys. Rev. Lett.* **2010**, *104*, 185702.

(67) Shulenburger, L.; Mattsson, T. R.; Desjarlais, M. P. **2015**, arXiv:1501.03850 [physics.chem–ph]. arXiv.org ePrint archive. `http://arxiv.org/abs/1501.03850` (accessed Sept 11, 2015).

(68) Snajdr, M.; Rothstein, S. M. *J. Chem. Phys.* **2000**, *112*, 4935–4941.

(69) Gálvez, F. J.; Buendıá, E.; Sarsa, A. *J. Chem. Phys.* **2001**, *115*, 1166–1171.

(70) Frisch, M. J. et al. Gaussian09 Revision E.01. Gaussian Inc. Wallingford CT 2009.

(71) NIST Database of basic atomic spectroscopic data; (accessed Aug 3, 2015). `http://physics.nist.gov/PhysRefData/Handbook/element_name.htm`, 2015.

(72) Paier, J.; Marsman, M.; Hummer, K.; Kresse, G.; Gerber, I. C.; Ángyán, J. G. *J. Chem. Phys.* **2006**, *124*, 154709.

(73) Paier, J.; Marsman, M.; Kresse, G. *J. Chem. Phys.* **2007**, *127*, 024103.

(74) Stroppa, A.; Kresse, G. *New J. Phys.* **2008**, *10*, 063020.

(75) Gao, W.; Abtew, T. A.; Cai, T.; Sun, Y. Y.; Zhang, S. B.; Zhang, P. **2015**, arXiv:1504.06259 [cond–mat.mtrl–sci]. arXiv.org ePrint archive. `http://arxiv.org/abs/1504.06259` (accessed Sept 24, 2015).

(76) Díaz, C.; Pijper, E.; Olsen, R. A.; Busnengo, H. F.; Auerbach, D. J.; Kroes, G. J. *Science* **2009**, *326*, 832–834.

(77) Wijzenbroek, M.; Kroes, G. J. *J. Chem. Phys.* **2014**, *140*, 084702.

(78) Libisch, F.; Huang, C.; Liao, P.; Pavone, M.; Carter, E. A. *Phys. Rev. Lett.* **2012**, *109*, 198303.

(79) Zhu, M.; Mitas, L. *Chem. Phys. Lett.* **2013**, *572*, 136–140.

(80) Tipton, W. W.; Drummond, N. D.; Hennig, R. G. *Phys. Rev. B* **2014**, *90*, 125110.

(81) Hamann, D. R. *Phys. Rev. B* **2013**, *88*, 085117.

(82) Vanderbilt, D. *Phys. Rev. B* **1990**, *41*, 7892–7895.

(83) Deible, M. J.; Kessler, M.; Gasperich, K. E.; Jordan, K. D. *J. Chem. Phys.* **2015**, *143*, 084116.

(84) Zimmerman, P. M.; Toulouse, J.; Zhang, Z.; Musgrave, C. B.; Umrigar, C. J. *J. Chem. Phys.* **2009**, *131*, 124103.

(85) Bouabça, T.; Amor, N. B.; Maynau, D.; Caffarel, M. *J. Chem. Phys.* **2009**, *130*, 114107.

(86) Al-Saidi, W. A.; Umrigar, C. J. *J. Chem. Phys.* **2008**, *128*, 154324.

(87) Schautz, F.; Buda, F.; Filippi, C. *J. Chem. Phys.* **2004**, *121*, 5836–5844.

(88) Abbasnejad, M.; Shojaee, E.; Mohammadizadeh, M. R.; Alaei, M.; Maezono, R. *Appl. Phys. Lett.* **2012**, *100*, 261902.

Graphical TOC Entry